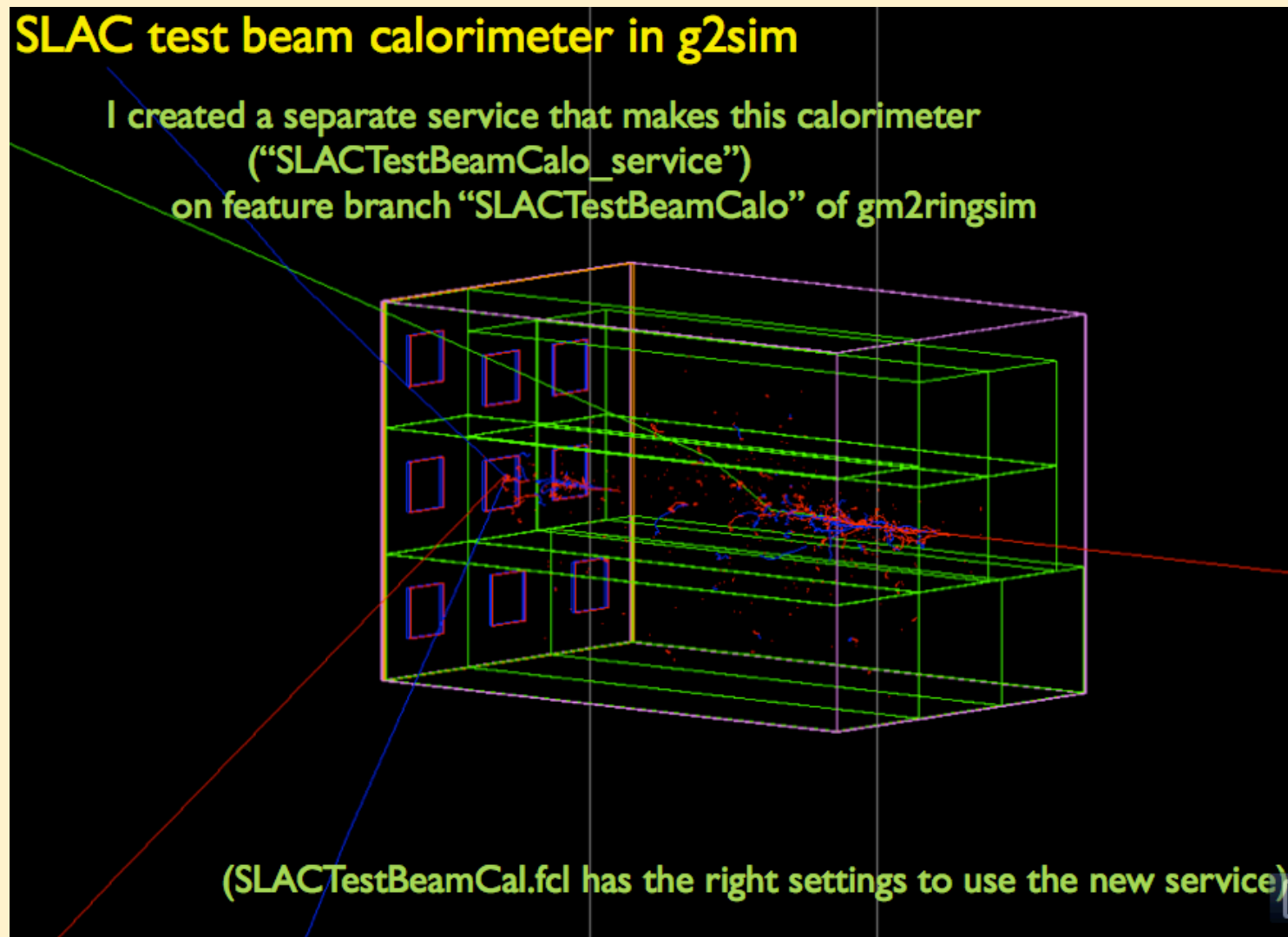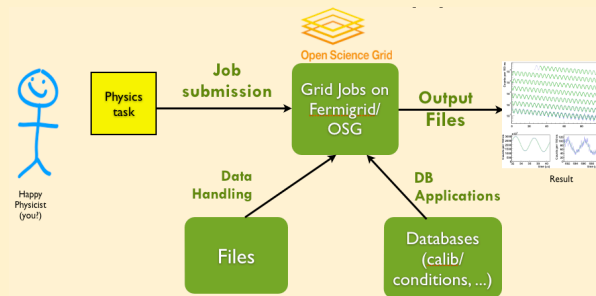# Simulations, Offline & Computing
(not necessarily in that order :-)
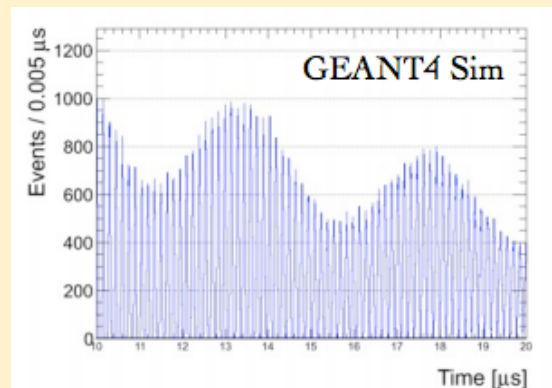
Adam Lyon, Fermilab Scientific Computing Division
December 2013 Collaboration Meeting

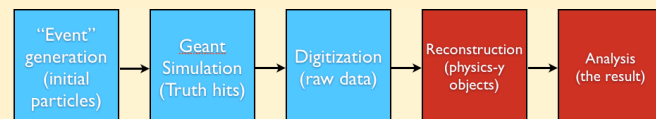# Since we've spoken last in April…

Fermilab computing update
(Important SCD reorg, new resources for experiments)

GEANT4 Sim

Simulation Progress

Evolution into Offline computing

Evolution into a mature computing environment

Significant Reorganization of the Scientific Computing Division

SCD is computing for science; CCD is the lab's IT organization

SCD reorganizes every 3-4 years. We did it again in June...

o Had several large departments – some were too big to understand their scope & mission

o Commitment to science was not immediately obvious

o "Service based" organization was not immediately obvious

o The huge CMS computing effort lived in its own world

The new structure – big departments become big quadrants each with smaller focused departments

Scientific Computing
Robert M Roser — Division Head

Lothar Bauerdick — Deputy Division Head
Ruth Pordes — Associate Division Head
Robert Tschirhart — Special Assignment Project X

Scientific Technical Architecture
(Ruth Pordes) — Lead

(Lothar Bauerdick)
Keith Chadwick
(Ian Fisk)

← **Commitment to common solutions**

Entrepreneurial Ventures
(Rob Roser) — Lead

(Gabriele Garzoglio)
Anne E Heavey

(Ruth Pordes)
(Panagiotis Spentzouris)

← **Collaborations and projects to advance computing and services**

Systems for Scientific Applications	Scientific Programs	Scientific Computing Services	Scientific Computing Facilities
Panagiotis Spentzouris — Associate Head	(Lothar Bauerdick) — Associate Head	Margaret E Votava — Associate Head	Stuart C Fuess — Associate Head

Systems for Scientific Applications
- o Scientific Software Infrastructure (art, CMSSW)
- o Simulations
- o Electronic Systems and Engineering
- o Real time Software

Scientific Programs
- o All scientists matrixed into in this quadrant: CMS, Theory, IF, Astro
  Science is an explicit activity of the SCD

Note close incorporation of CMS

I head this department and am deputy quadrant head of SCS

Scientific Computing Services
- o Grid and Cloud development and operations (+ Tevatron Data Preservation)

→ o **Scientific Data Processing development and operations**
  – Data handling and management
  – Database applications
  – CMS processing operations

Scientific Computing Facilities
- o High Performance and Parallel Computing (e.g. LQCD)
- o U.S. CMS Computing Facilities
- o Data Movement and Storage
- o Fermilab Experiments Facilities

Service offerings are clearer

Not listed here – Liaisons

You yawn: Why do we care?

o We benefit from commitment to common solutions (Data management, Grid and cloud computing, sysadmins)

o We benefit from closer collaboration with CMS
  – Tap into their extensive computing experience
  – Beefs up human resources for Intensity Frontier projects
  – While CMS and IF solutions will most likely remain separate for now, CMS & IF computing will evolve with common principles in mind
  – art may eventually remerge with CMSSW [good for us and art's future]

o We benefit from a service organization
  Part of ISO2K (standards for IT) and ITIL (best practices for IT)
  e.g. The Computing Sector does not operate by the seat of its pants
  Offerings, responsibilities, and expectations are clear [takes a lot of work to do right]

Upcoming bureaucracy

January 2014:
FY14 SPPM (Scientific Projects Portfolio Management) Review
Determine the yearly strategy for purchases that will be on the floor early 2015

Later in 2014:
Muon g-2 Computing TSW – A big deal for the SCD
Couples the experiment to SCD Service Level Agreements

Tangible Improvements for you

o **Obtaining computing resources for new collaborators is vastly improved (result of the 2012 S&T review)**

o **You fill out the standard visitors badge form (or similar)**

o **That automatically triggers**
  - **A request for approval by the spokes**
  - **Your user ID granted**
  - **Kerberos and service passwords created**
  - **A home area created in AFS**
  - **An account for you in Redmine**
  - **A grid certificate within the Fermilab VO and g-2 group for jobs**
  - **A task for us (Brendan K and me) to grant you g-2 specific resources (gm2gpvm account, Redmine groups, mailing lists)**
  - **All of this is tracked by the service desk (easy to spot when something goes wrong)**

Muon g-2 welcome page

Welcome to Muon g-2

Welcome to Muon g-2
    Obtaining g-2 computer accounts
    Using Redmine
    General documentation
    Information about Computing Resources
    Mailing lists
    If you need help

Congratulations on joining the g-2 experiment! This welcome packet has information on how to get computing accounts and what to do when you've got them. PLEASE READ THIS PAGE CAREFULLY as many common questions are answered here.

Obtaining g-2 computer accounts

(If you already have an account on the gm2gpvm machines, you can skip this section. If you don't know what gm2gpvm is, then execute this section).

1. If you don't have a visitor's badge, your badge has expired, or have never had any kind of Fermilab computer account, you need to carefully follow ⊡ these instructions . BUT READ THESE NOTES FIRST:

- Applying/Renewing/Obtaining a visitor ID, badge, and computing accounts are all the same process
- You may want to consider yourself an "Off-site visitor" for the purposes of filling out forms
- You will have to choose your "affiliation" from a drop down list. Be sure to choose "E-989 Muon g-2". Do **NOT** choose your university.
- Wait for your kerberos password. Then wait some more for the "account creation" process to complete. Once this is done, you should have an account on gm2gpvm, Redmine and git access.
- When you receive your Kerberos Password, come back here to continue the process of getting Muon g-2 resources.

2. If you don't remember your Kerberos Password AND/OR you need a Services password (needed to update this Wiki) then simply call the Fermilab Service Desk at 630-840-2345 and tell them you need **both** of those passwords. If you are at Fermilab, you can go in person to the Service Desk on the ground floor of Wilson Hall across from the Argonne Credit Union branch. In my experience, the people who man the service desk are friendly and helpful. If you have a different experience, be sure to let Adam know at lyon@fnal.gov.

3. If you already have Fermilab computing access from a different experiment (and you already have a visitors or employee badge), but you don't have access to g-2 computing resources, then do the following,

- Go to ⊡ https://fermi.service-now.com
- If the "Self Service" page doesn't come up, look at the menu bar on the left side of the screen and select "Self Service".
- On the "Self Service" page, under "Request Things" select "Request New / Renew Accounts"
- On the page that appears, select "Affiliation/Experiment Computing Account Request"
- Fill out the page, being sure to select "E989 Muon g-2" as the new Affiliation/Experiment (do **not** choose your university if it is listed)
- This will start the process of getting you accounts. Wait for it to complete (you'll get a service desk ticket assigned)

4. If you have any other problems or questions, don't hesitate to send mail to gm2-computing@fnal.gov .

Running jobs on remote grid sites

NOvA has pushed a full end-to-end test of running on the Open Science Grid

o Important since Fermilab will not have enough resources for everyone
o Important for contributions of computing resources

Needed:
– Portable software distribution via CVMFS
– Data management with SAM
– Art compatible with data management
– Cache for data at Fermilab and OSG sites
– Working job submission system
– Mechanism for returning output data back to Fermilab (still fragile)
– Properly configured and working remote sites (often the hardest part)
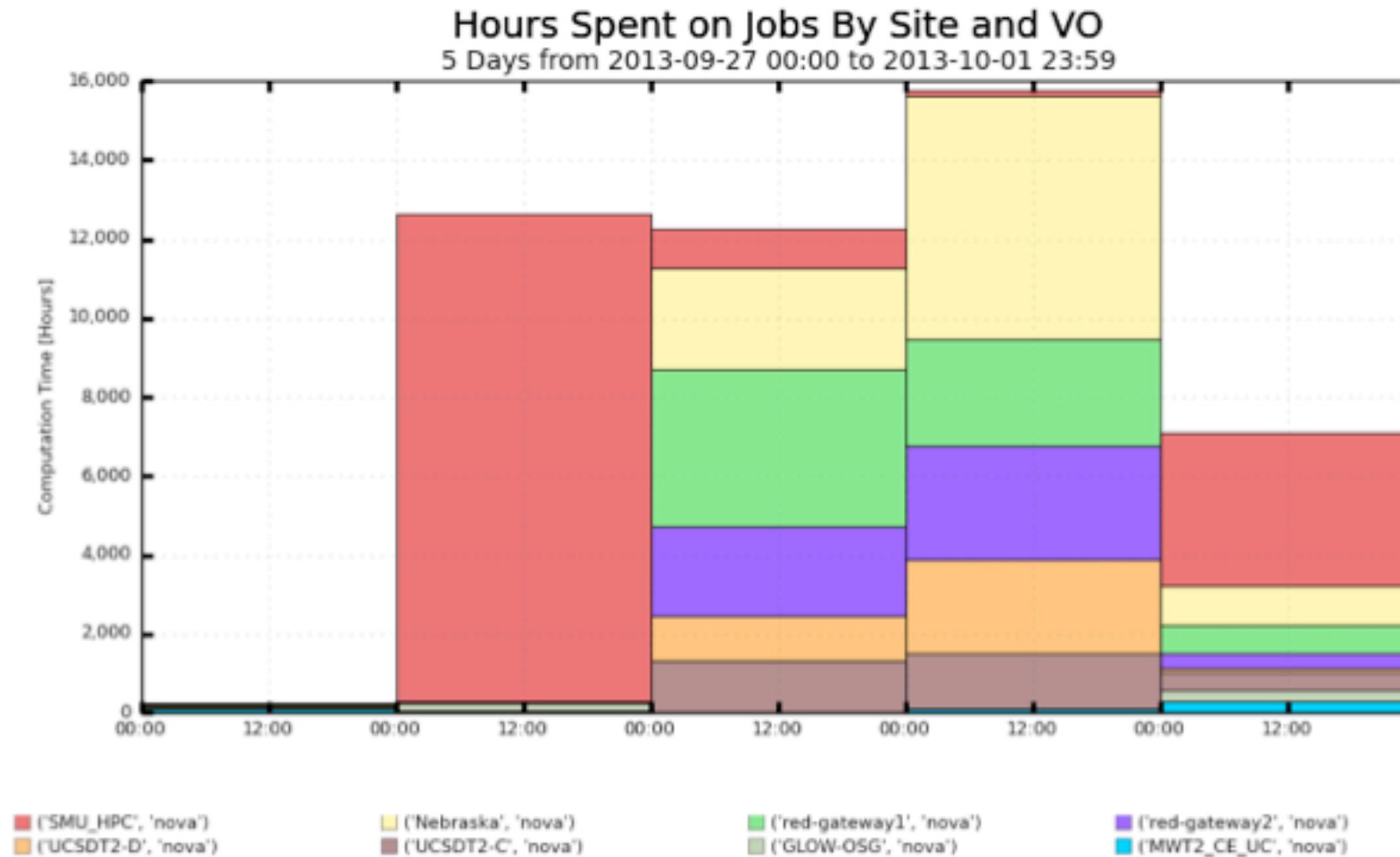
NOvA pushed this and made it a big success
Microboone is next
We should follow close behind (if you have a CMS or ATLAS Tier2, Tier3 or generic OSG site, you can contribute computing resources to the experiment)

Running jobs on remote grid sites


Use Case: OSG Processing

Hours Spent on Jobs By Site and VO
5 Days from 2013-09-27 00:00 to 2013-10-01 23:59

Maximum: 15,758 Hours, Minimum: 247.98 Hours, Average: 9,602 Hours, Current: 7,106 Hours

Legend:
- ('SMU_HPC', 'nova')
- ('UCSDT2-D', 'nova')
- ('Nebraska', 'nova')
- ('UCSDT2-C', 'nova')
- ('red-gateway1', 'nova')
- ('GLOW-OSG', 'nova')
- ('red-gateway2', 'nova')
- ('MWT2_CE_UC', 'nova')

Red	= SMU	= NOvA Institution
Yellow+Green+Purple	= U. Nebraska Lincoln	= General OSG site
Orange+Brown	= UCSD	= General OSG site
Light Bluew	= U. Chicago	= Atlas Tier 2
Grey	= U. Wisc	= CMS Farm

Notes:
- Global set is 9740 fcl files worth of cosmics

Busy time per file
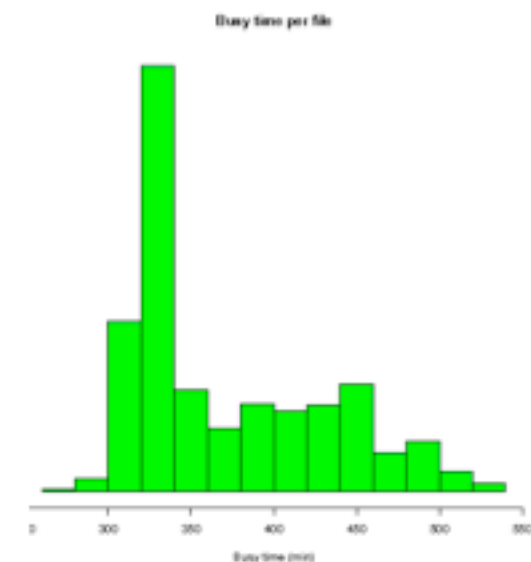
- Average Run Time is ~330 minutes (5.5 hrs for 100 events)

A big cache

Recently purchased 4 PB of disk cache for IF experiments

Cache: Temporary non-critical storage of files for data processing
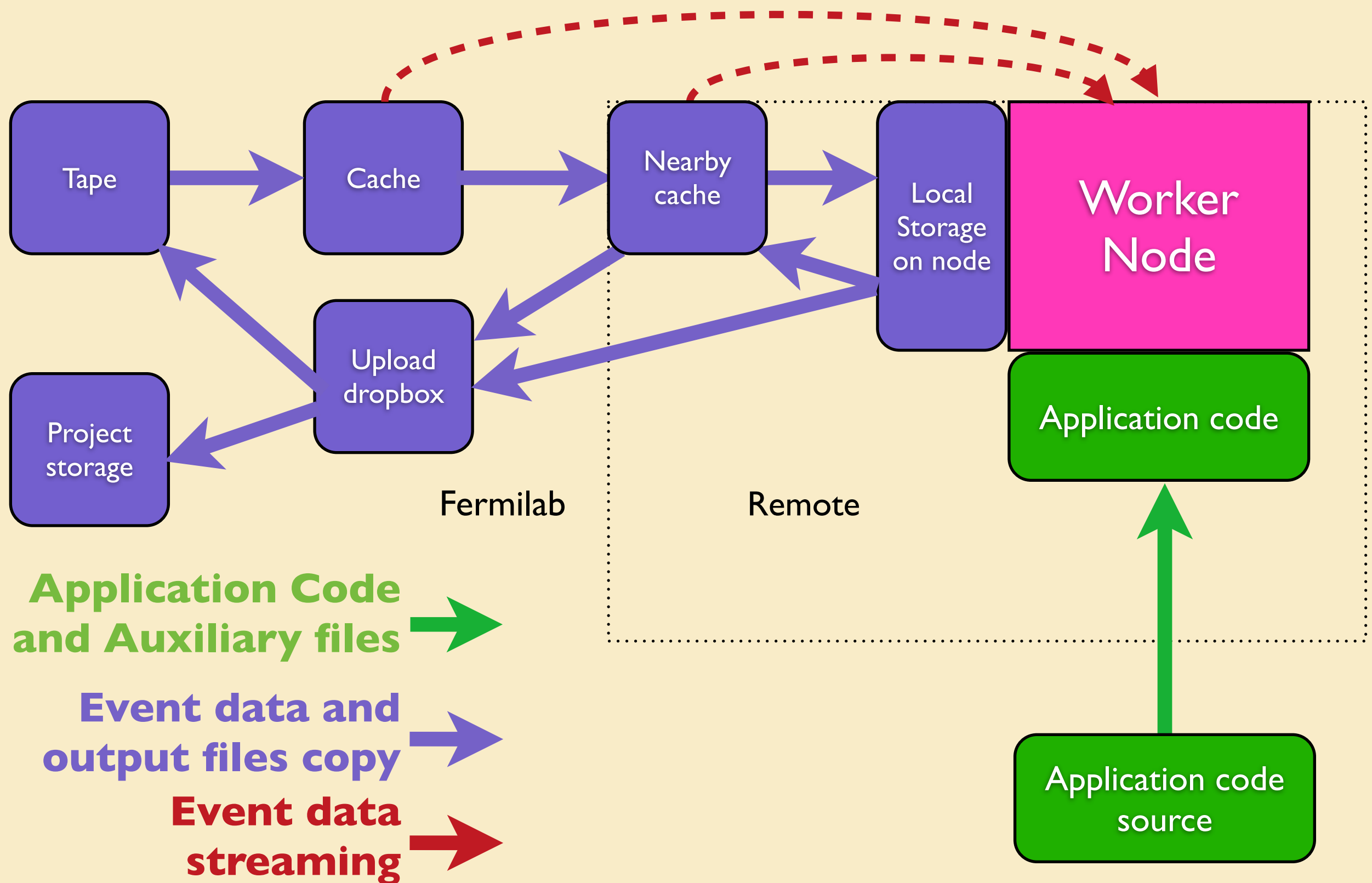
Managed by caching system (dCache) and supervised by data management system (SAM)

Current plan: Allocate a small "write pool" per experiment for storing online data to tape

Allocate 2 PB of "read pool" for data cached from tape

Future: "read-write pool" for non-tape backed up data

Data Handling Functionality

Tape → Cache → Nearby cache → Local Storage on node → Worker Node

Project storage

Upload dropbox

Application code

Fermilab | Remote

Application code source

Application Code and Auxiliary files →

Event data and output files copy →

Event data streaming →

FabrIc for Frontier Experiments

A collection of projects that provide <u>common</u> computing services and interfaces needed to turn a physics task into results

Open Science Grid

Physics task

Job submission

Grid Jobs on Fermigrid/ OSG

Output Files

Counts per 150 ns

Result

Happy Physicist (you?)

Data Handling

DB Applications

Files

Databases (calib/ conditions, ...)
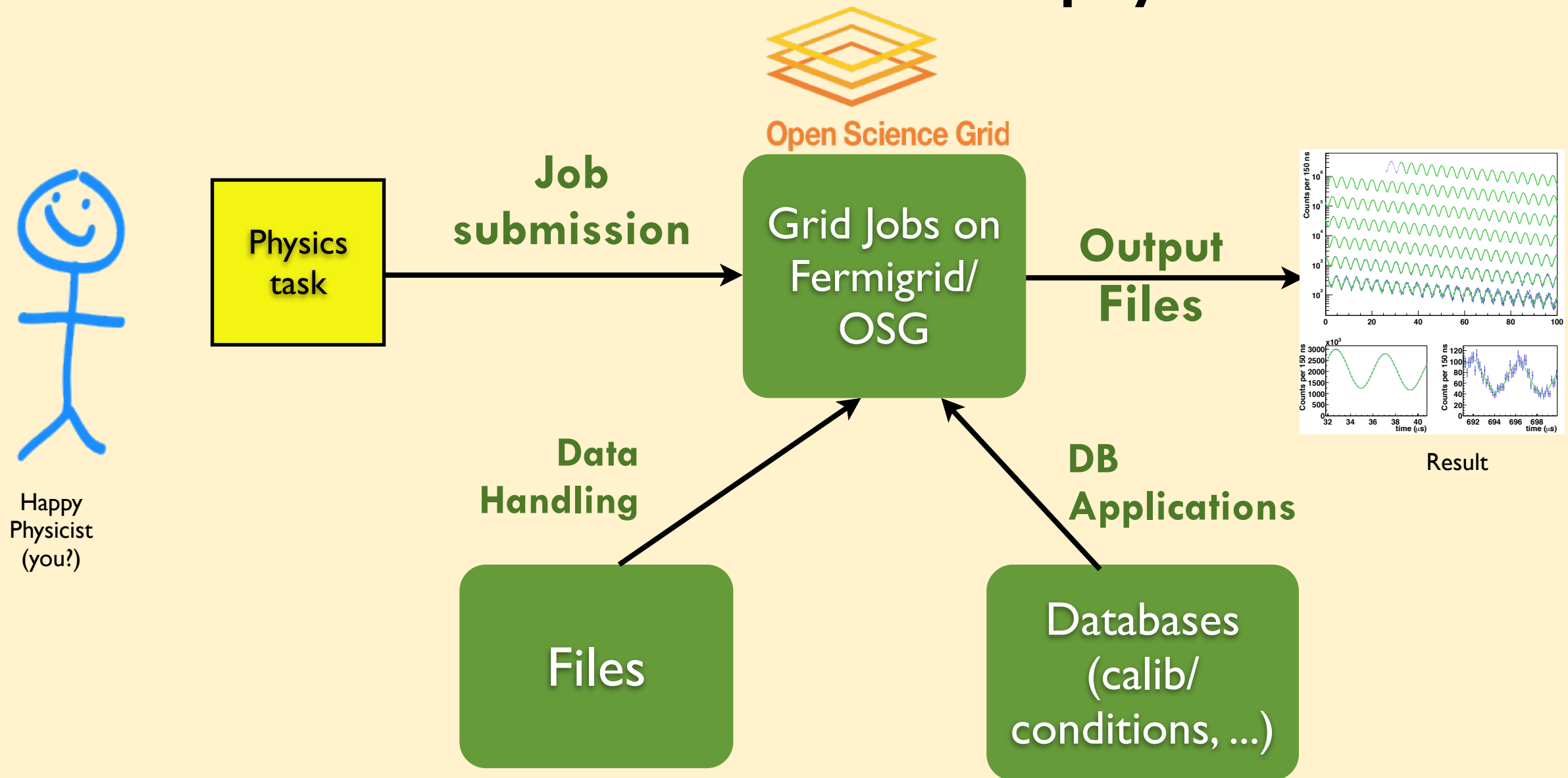
Tailored for opportunistic running

art Workbook & Users Guide

Training for learning the art framework and C++ best practices

– See **https://cdcvs.fnal.gov/redmine/projects/g-2/wiki/ArtWorkbook**

o **Many examples to learn art concepts**
o **Work on a toy collider detector**
o **Basic C++ knowledge**
o **300 pages and not done yet!**

**There will be C++ classes (from a physicist) in the summer**

On to g-2 stuff...

End-to-End Beam Simulations

See J-F Ostiguy's talk at Beam/Ring Breakout yesterday

Muon g-2
Beam lines

• An 8.89 GeV/c proton bunch, 120 ns long, is transported to the Target Station via AP-1 at an average rate of 15 Hz, with 100 Hz bursts (20 bunches, 10 ms interval)
• A 3.1 GeV/c Positive secondary beam travels down AP-3 and is injected into the Debuncher in the 30 straight section with Lambertsons and a kicker
• Some of the pions decay into 3.09 GeV/c muons as they travel down AP-3
• Muons can circle the 550 meter Debuncher as many times as desired
• The abort located in the 50 straight section can be used to remove protons
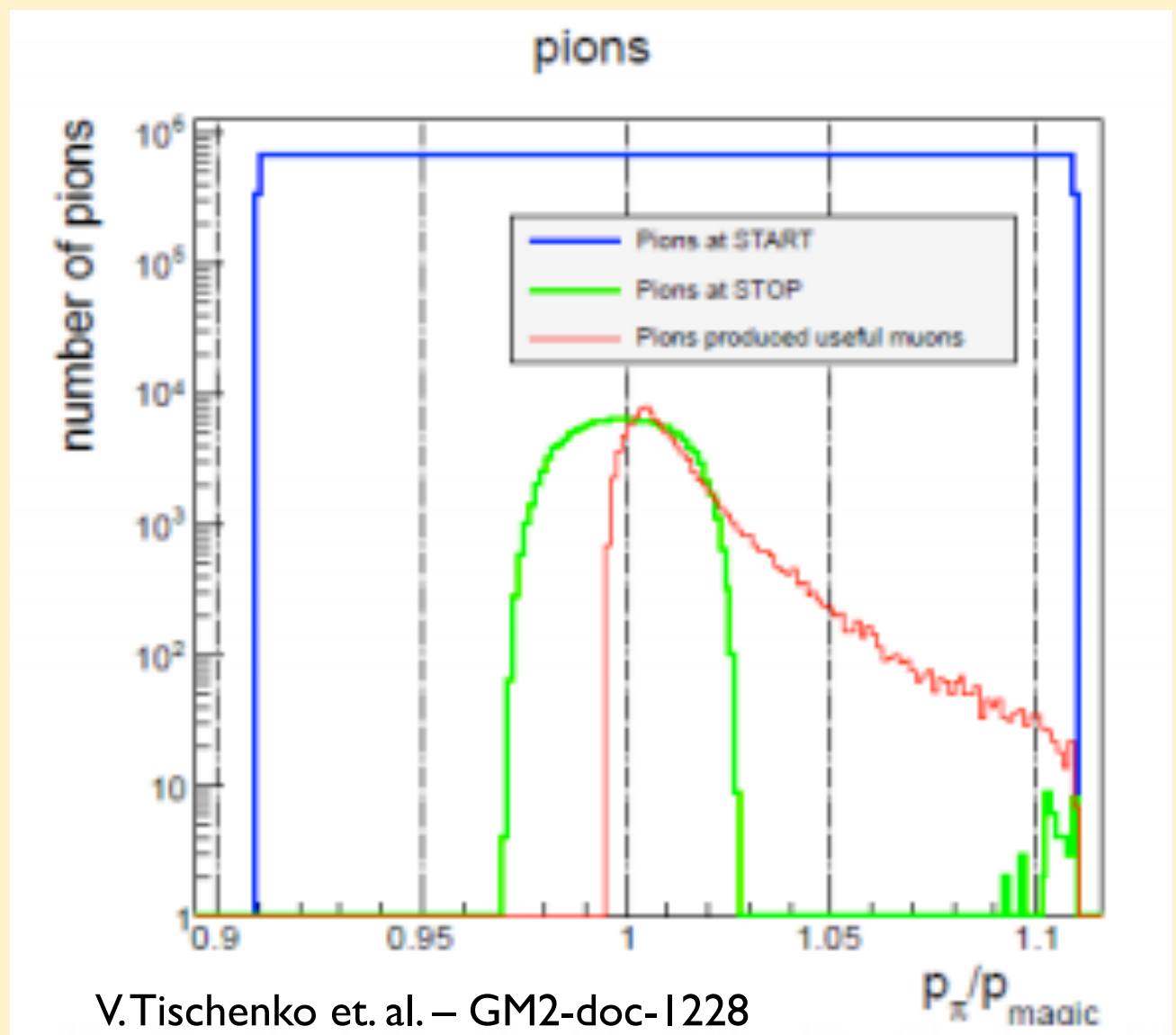• 3.09 GeV/c muons are extracted into a beam line that transports them to the experiment
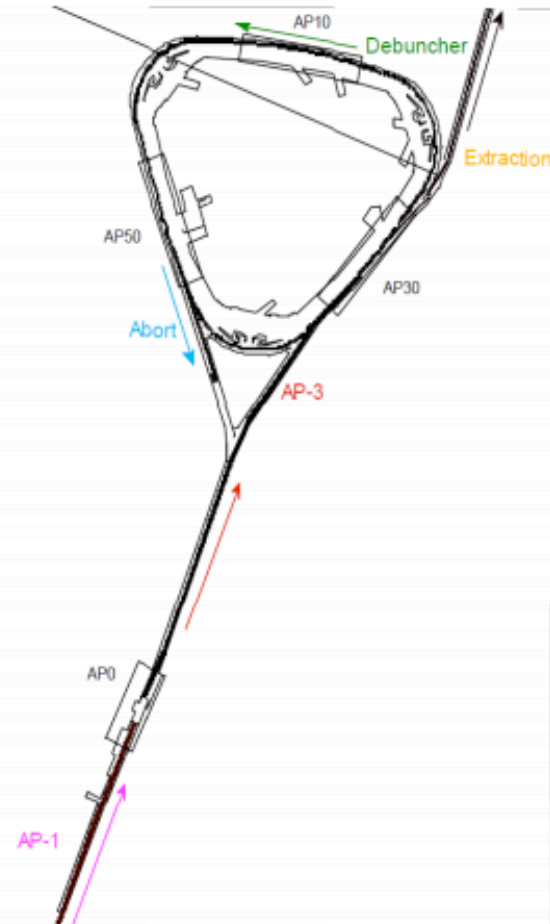
Jim Morgan doc-db-148

In January expect to have useful distributions of muons entering g-2 ring

Target to the g-2 Ring

Uses G4BL (many improvements over MAD8) And MARS for pion production in target


V. Tischenko et. al. – GM2-doc-1228

Ring Simulation

o **The port of g2migtrace to art started about a year ago**
o **It's been complete for many months and is in heavy use**
o **Now used by a group of mostly students and postdocs**
o **I'm amazed by what people are doing with this system**

o **Simulation of SLAC test beam calorimeter while maintaining sanity (YAY!) [Robin]**
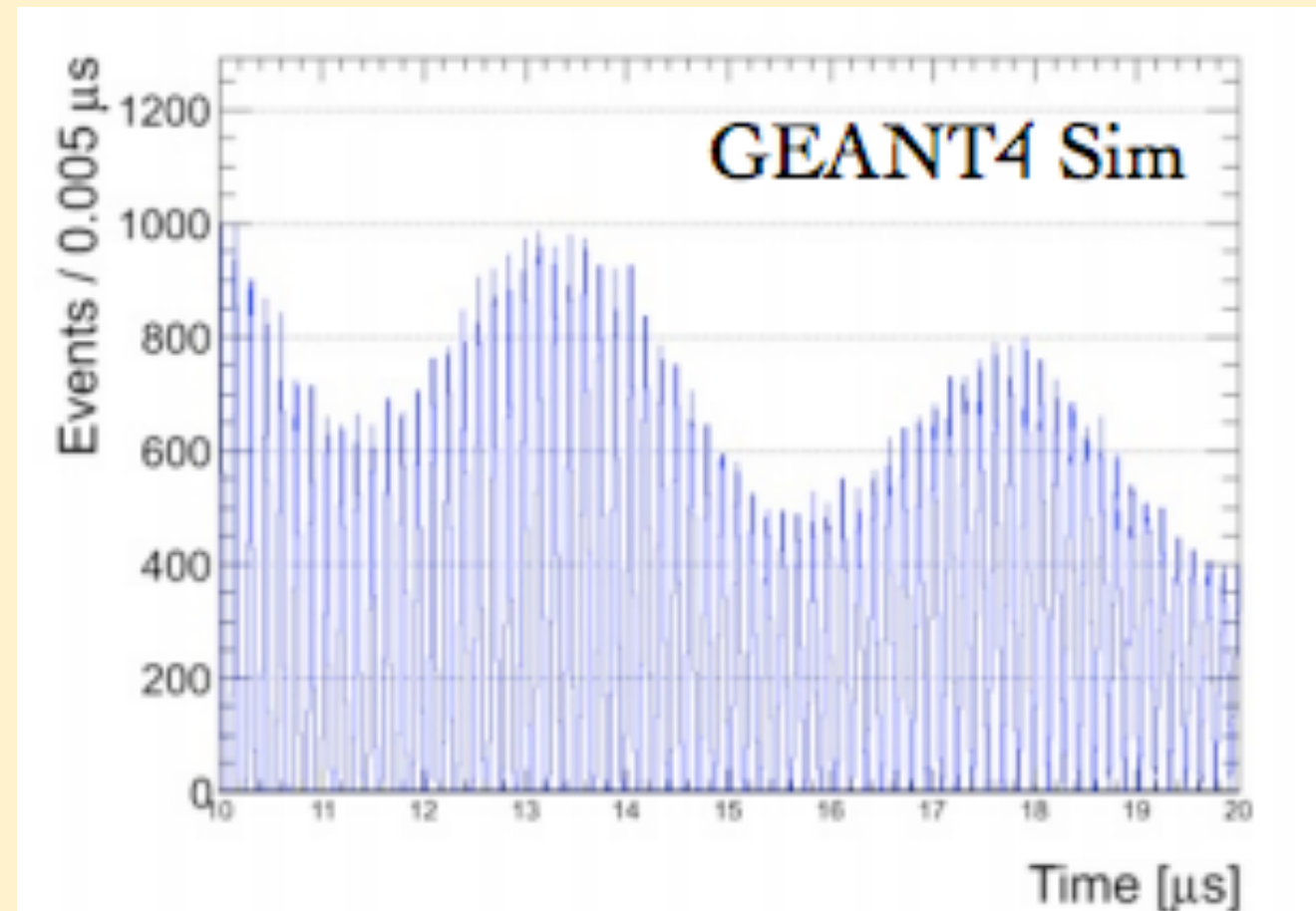

SLAC test beam calorimeter in g2sim

I created a separate service that makes this calorimeter ("SLACTestBeamCalo_service") on feature branch "SLACTestBeamCalo" of gm2ringsim
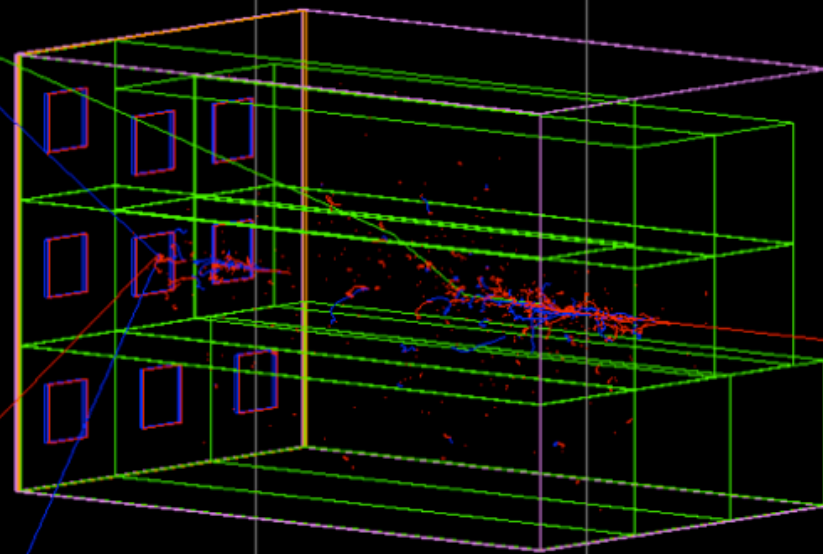
(SLACTestBeamCal.fcl has the right settings to use the new service)

Fast Rotation/CBO Studies

Amazing studies by Thomas Gadfort

- Start a pencil beam with $\delta P/P = 0.5\%$ and let it run for 30 μs.
- Compute the angular distance from the center of the muon bunch vs time.
  - i.e., what fraction of the ring is filled with the muon bunch.

Delta function bunch at $t=t_0$

After ≈ 24 μs the bunch starts to lap itself.

GEANT4 Sim
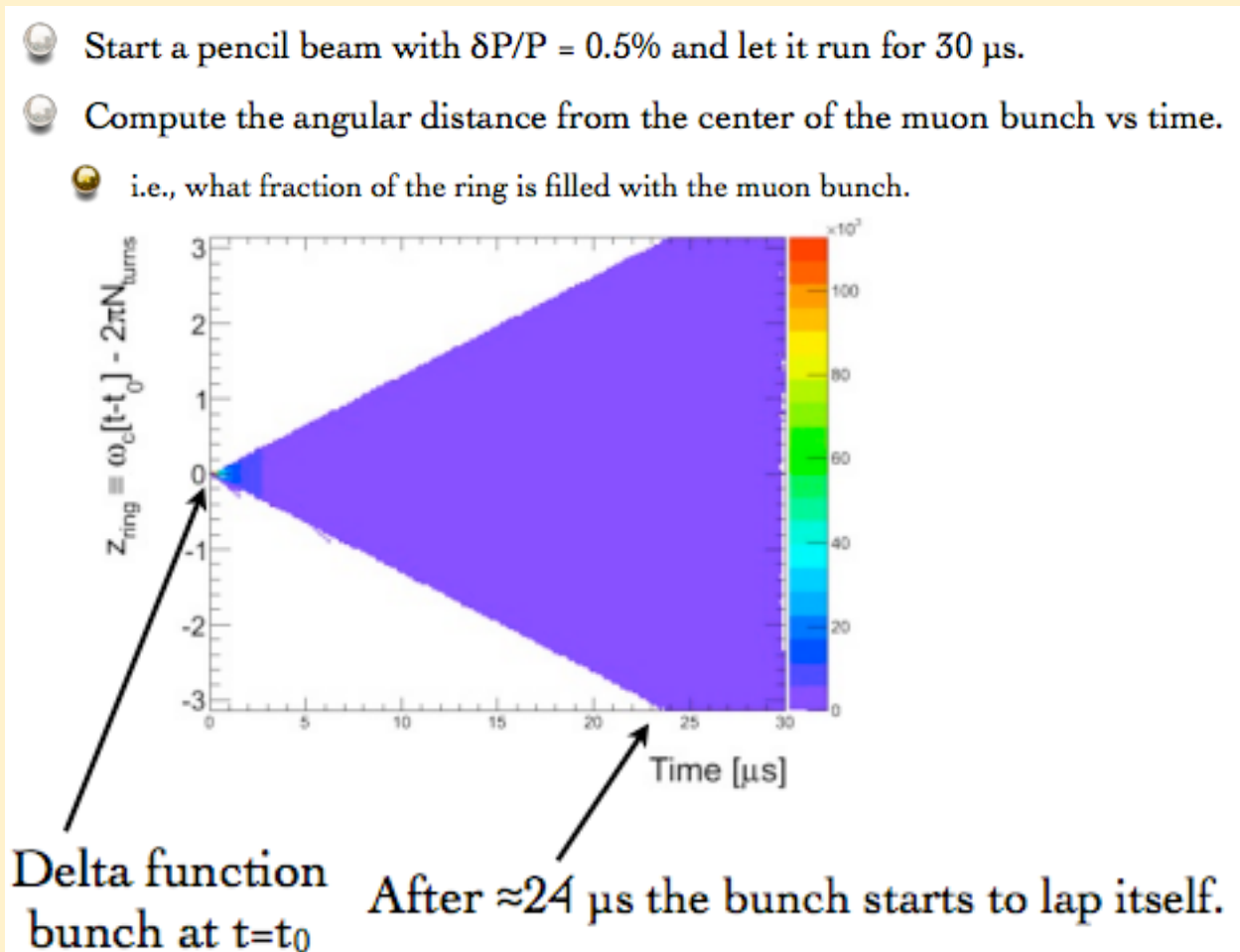
Very fast Muon "gun"

o **Just need to know magnetic fields and initial muon parameters and can *calculate* where muon will be at any time**
o **Includes CBO**
o **Can generate 100 muons/s/batch slot – 10M in a few hours is possible**
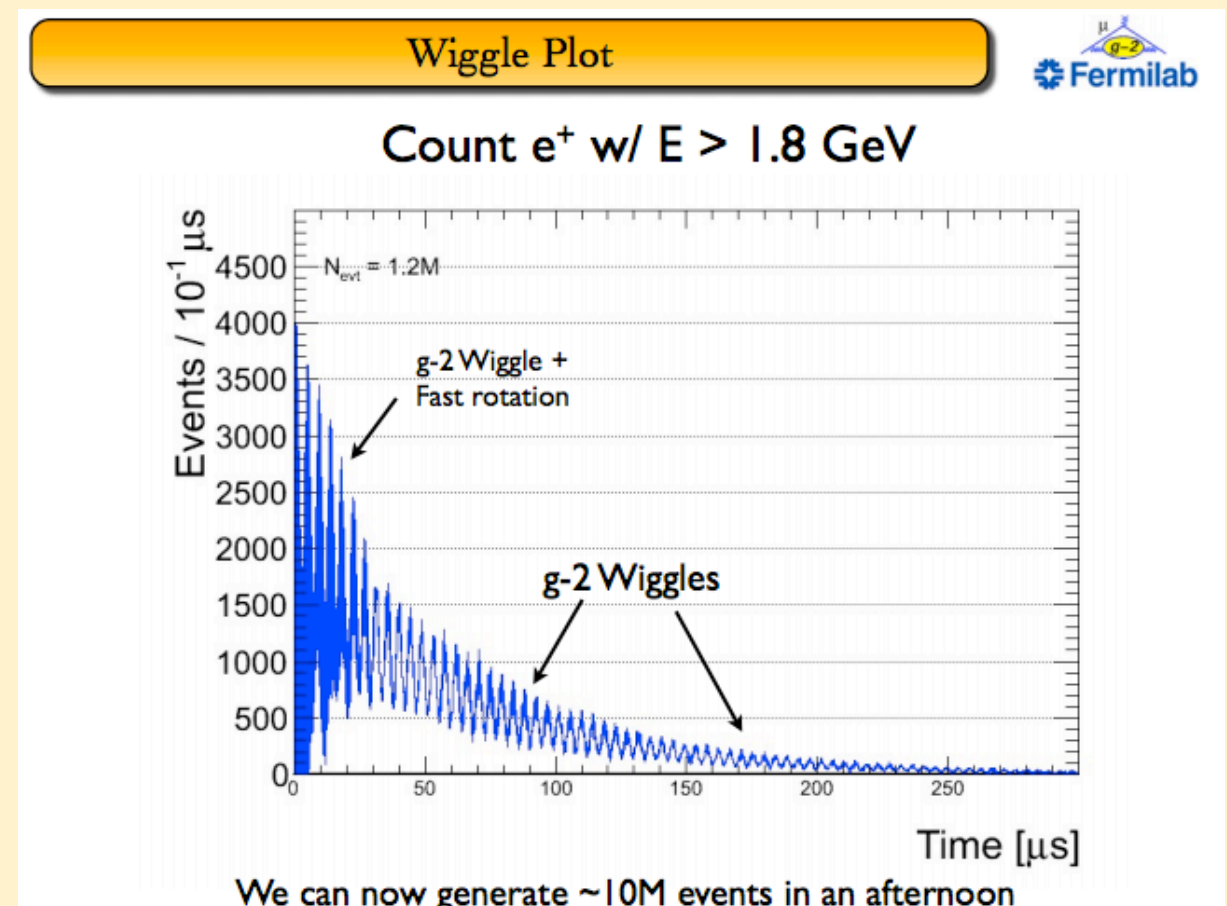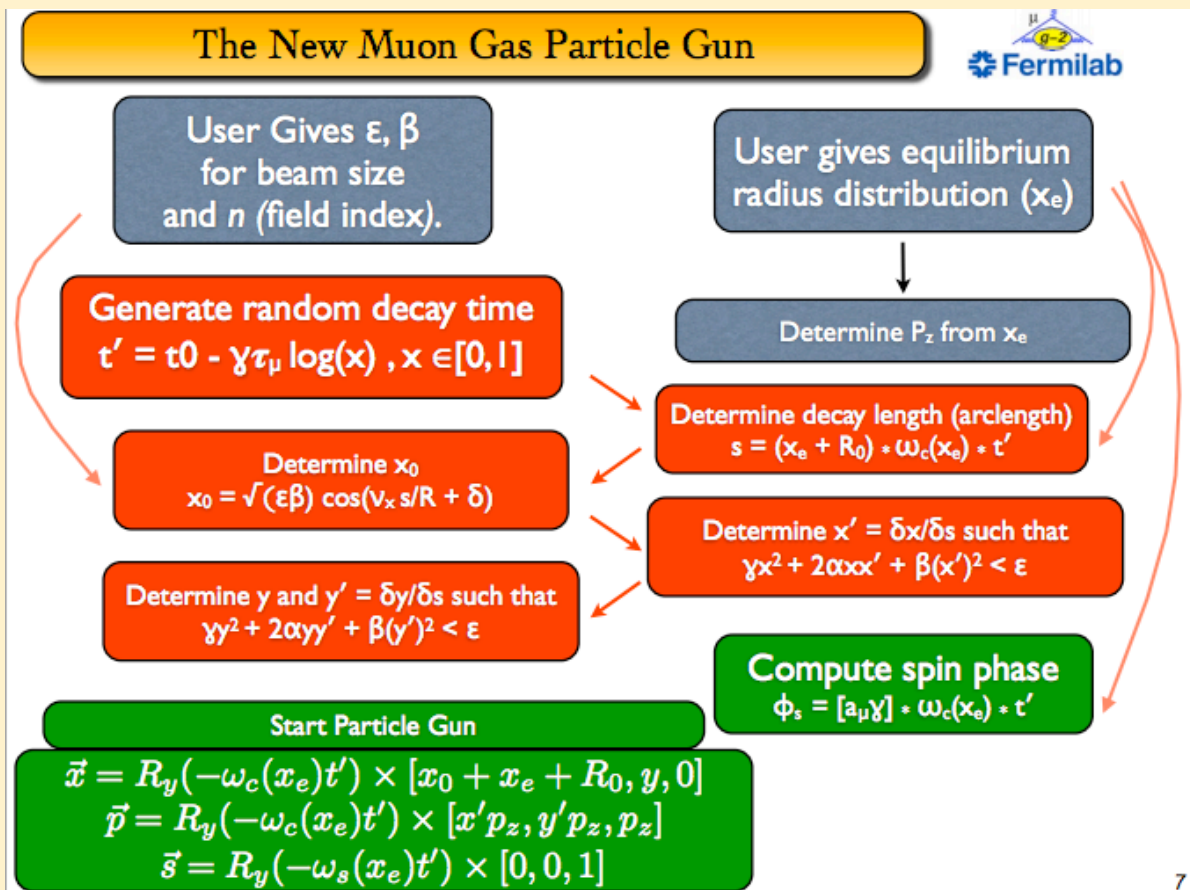
See Thomas' Beam/Ring talk and plenary Friday talk

Tracker material and configuration

Simulates fit of trackers in scallop
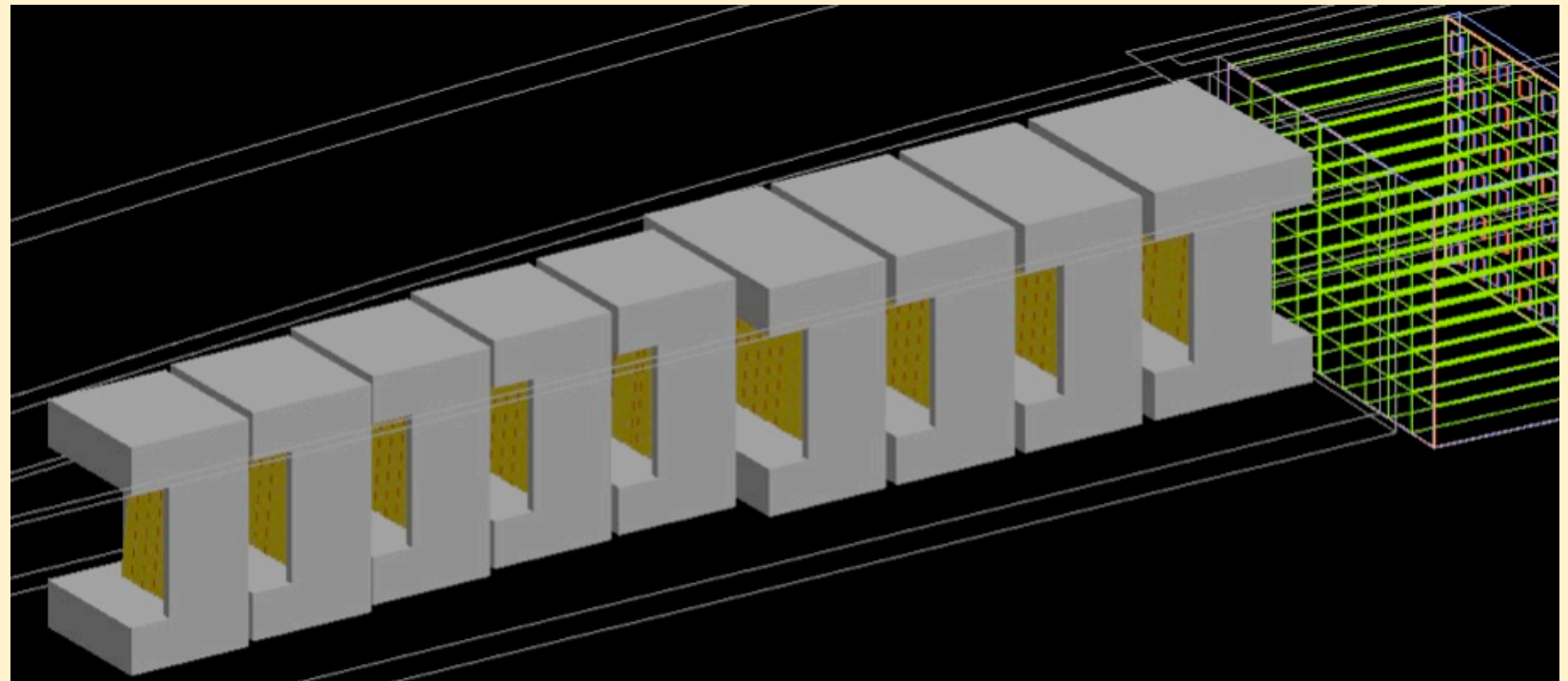
Simulate effect of different materials in tracker on positron energy in calorimeters

Positron Energy Over All Calorimeters

Legend:
- w/ tracker 1.5mm Thick Manifold
- w/tracker 3mm Thick Manifold
- w/tracker 0.1mm Thick Manifold
- w/tracker [straw material only]
- w/o tracker

Simulation challenges

We are writing the tools necessary for important studies

o (Hertzog) Early-to-late Gain Stability on long and short timescales within a fill

Determine how sensitive we must be in both hardware and offline analysis to keep the systematic under control
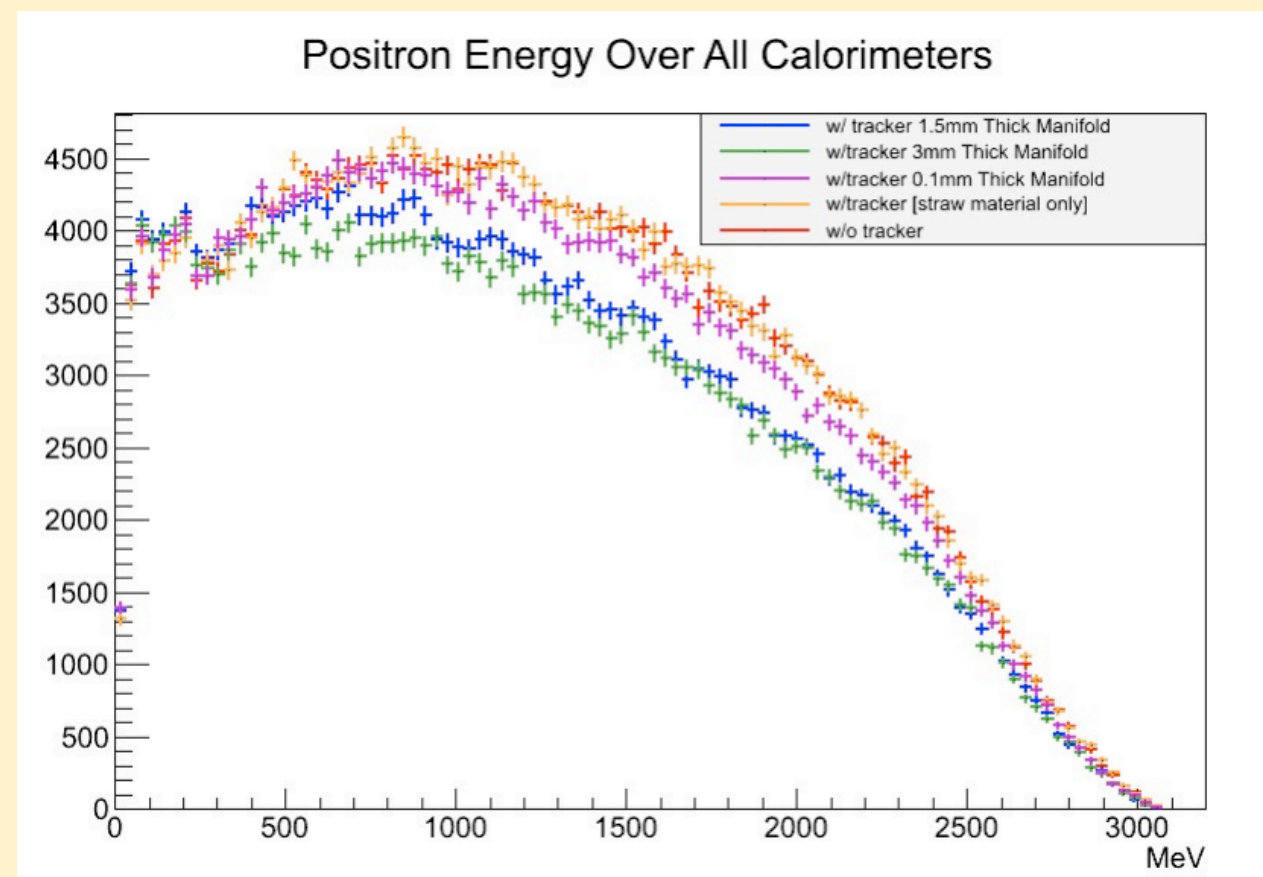
Requires a full omega_a analysis

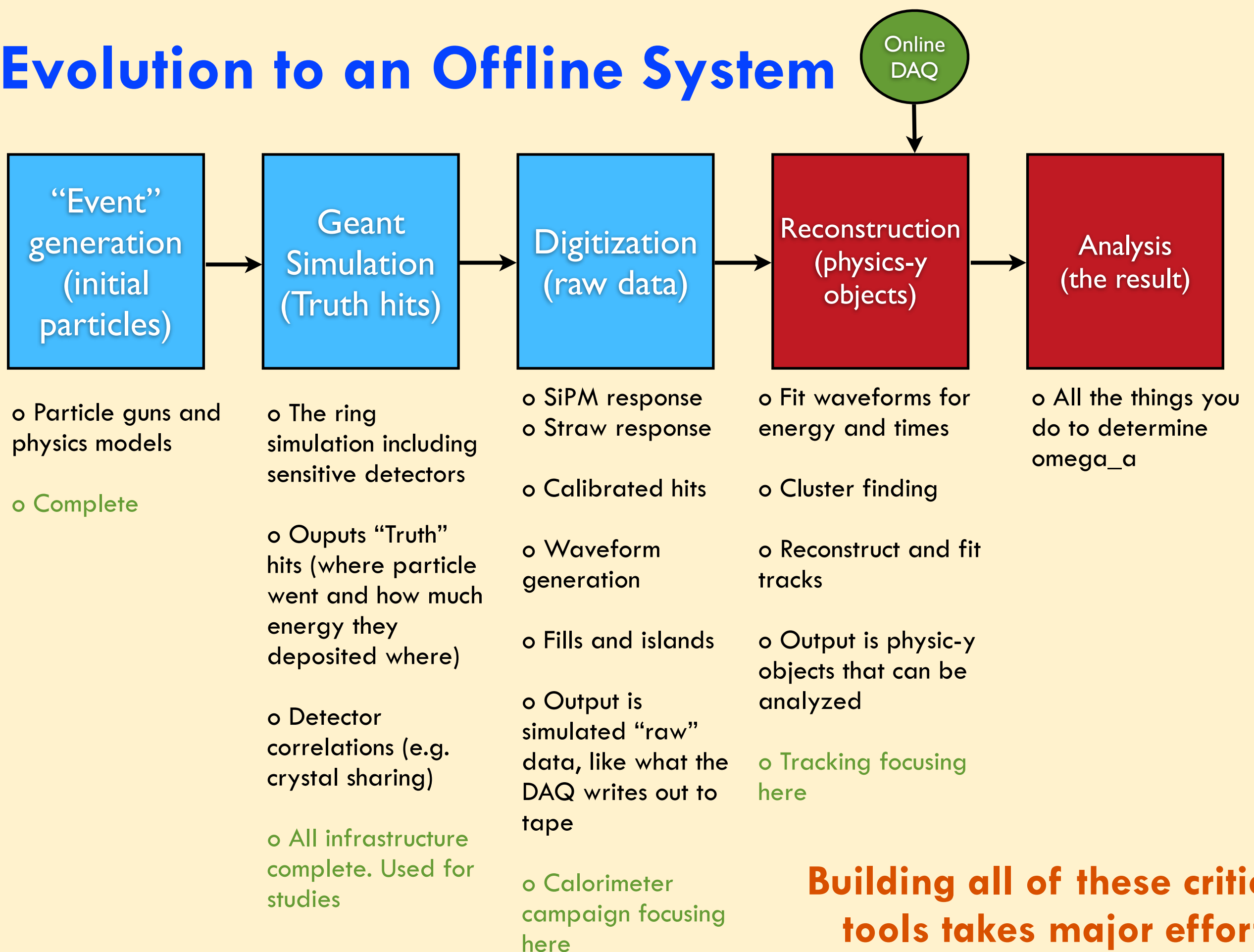**o (Casey) Tracking simulation challenges**
- **Pileup & Early-to-late calibration**
- **Understanding the fit to the pileup subtracted time distribution**
- **Systematics of pitch correction**
- **E-field correction**
- **What is the value of a 2nd or 3rd tracker**
- **E/p with early-to-late systematics**
- **Collimator placement**
- **material budget**

Requires some sort of tracking code

Evolution to an Offline System

Online DAQ

"Event" generation (initial particles)	Geant Simulation (Truth hits)	Digitization (raw data)	Reconstruction (physics-y objects)	Analysis (the result)

o Particle guns and physics models

o Complete

o The ring simulation including sensitive detectors

o Ouputs "Truth" hits (where particle went and how much energy they deposited where)

o Detector correlations (e.g. crystal sharing)

o All infrastructure complete. Used for studies

o SiPM response
o Straw response

o Calibrated hits

o Waveform generation

o Fills and islands

o Output is simulated "raw" data, like what the DAQ writes out to tape

o Calorimeter campaign focusing here

o Fit waveforms for energy and times

o Cluster finding

o Reconstruct and fit tracks

o Output is physic-y objects that can be analyzed

o Tracking focusing here

o All the things you do to determine omega_a

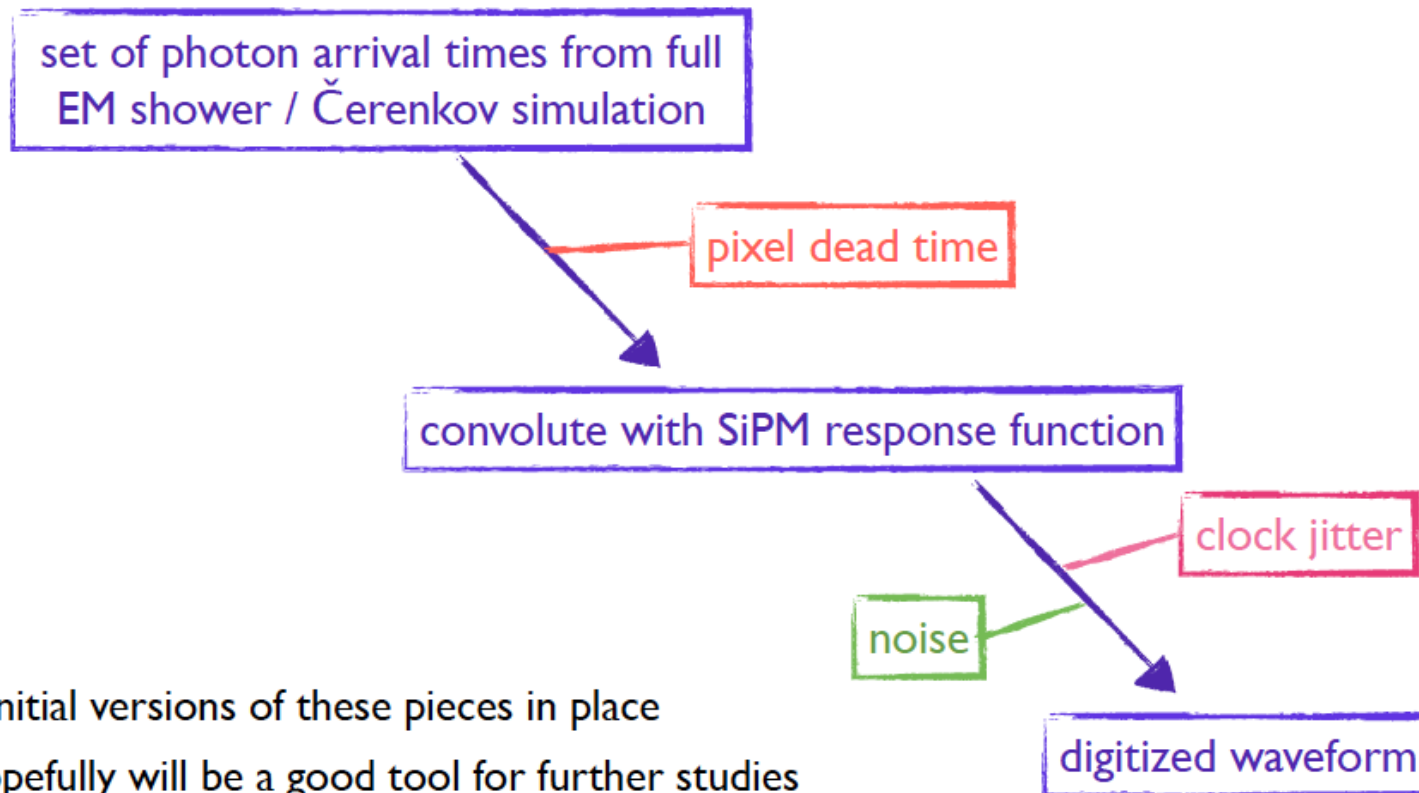**Building all of these critical tools takes major effort**

Calorimeter Campaign (UW+Cornell)

waveform simulations

Robin

I am working on a waveform module in art
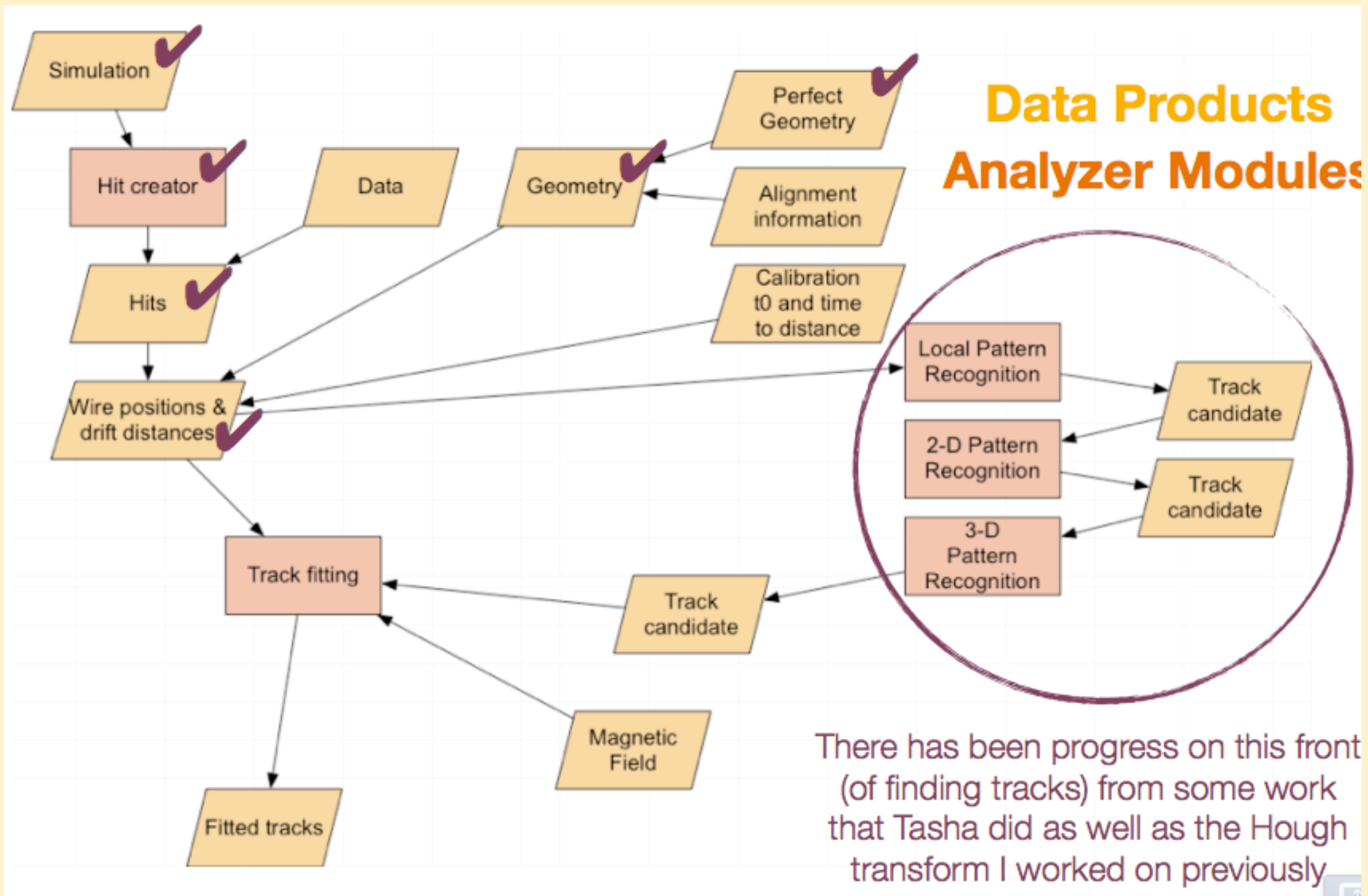photodetector photon hits from gm2ringsim ➔ digitized waveforms

set of photon arrival times from full EM shower / Čerenkov simulation

pixel dead time

convolute with SiPM response function

clock jitter

noise

I have initial versions of these pieces in place
hopefully will be a good tool for further studies

digitized waveform

Pete and others doing lots of work on waveform gymnastics (converting single muon events into fills, island chopping, etc).
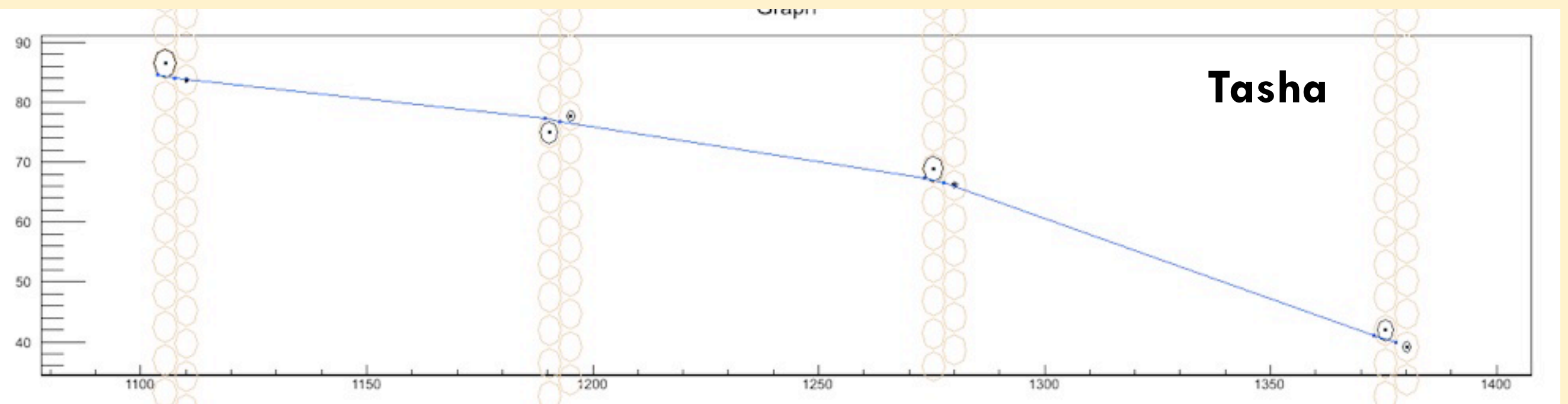
See **https://cdcvs.fnal.gov/redmine/projects/g-2/wiki/CalorimeterCampaign**
See Robin's talk next!

Tracking (Leah and NWU)

Data Products
Analyzer Modules

Simulation → Hit creator → Hits

Data

Perfect Geometry → Geometry ← Alignment information

Calibration t0 and time to distance

Wire positions & drift distances → Track fitting

Local Pattern Recognition → Track candidate

2-D Pattern Recognition → Track candidate

3-D Pattern Recognition

Track candidate

Magnetic Field

Fitted tracks

There has been progress on this front (of finding tracks) from some work that Tasha did as well as the Hough transform I worked on previously

Tracking

Tasha

Blue lines: Connecting Geant truth hits
Black dots: Center of hit straws
Black circles: drift circles

Determining t_0, trying drift functions, etc
Visualization!

See Leah's talk later this session

Evolution to a mature computing environment

o **Our code has many dependencies (art, Root, geant4, ...)**

   art updates about once a month sometimes with backwards incompatible changes

o **We sat at the same version of dependencies for a _year_**

   Not chasing latest and greatest served us well (allowed us to focus on the port)

o **But eventually, there are new features in dependencies that we want**

o **Last month I made a release with the latest and greatest of everything**
   – **Shipped via CVMFS (with OSG OASIS - if you have OSG you have it now)**
   – **Brendan K and I are solving backwards incompatibility problems**
   – **Will migrate over to this release in early January**
   – **Supports SLF5, Mac OSX 10.8 (Mountain Lion), SLF6 (soon)**

o **I'm working with LBNE on a new "gm2d" →"mrb"**

o **But this brings up lots of questions...**

Time for maturity

o **How often do we migrate to new dependencies?**

o **What release structure do we want?**
[Do you build all g-2 code yourself; do we make official and test releases; how often? Validation? Need a release manager?]

o **Additional infrastructure:**
Databases: Geometry, alignment, calibration

o **Can't wait too long to start on these things**

o **Need strong communication between tracker & calorimeter groups**

Workshops

October 2012 – Art & ArtG4 (this one got the ball rolling)
December 2012 – Art recap
March 2013 – Analyzing art output
July 2013 – Module and data products architecture
December 2013 – Taking stock and moving forward
                    [bad weather and other situations limited attendance]

Future workshops

January 2014 – Art, gm2sim, and offline introduction for new people
Early 2014 – Taking stock and moving forward redo

Doing these over ReadyTalk has worked well

Conclusions

o **Changes within the SCD are good for experiments**
o **Many resources and services are coming to fruition**

o **Enormous progress in simulations**
o **Evolving into a real offline system**
o **Need to complete the tools and get to studies**

o **I think the promise of art and our development environment (e.g. git) has been fulfilled**
  – **Allows us to work together :-)**
  – **Allows us to maintain our sanity :-)**
  – **Allows us to concentrate on physics and the experiment :-)**

 But things are getting more complicated. As we proceed on all fronts, we need to evolve our computing system into maturity too! And better documentation!!

Robin and Leah will talk about specifics in Calorimeter and Tracking

ADDITIONAL SLIDES

Service organization under ISO2K/ITIL

Various parts of the division offer *Services*
(e.g. e-mail, data management, batch slots)

A Computing TSW (Technical Scope of Work) defines the responsibilities and expectations between the experiment and the SCD Services. A living document

TSW's refer to SLA's (Service level agreements)
e.g. If a tape is unreadable during business hours, it will be pulled from service within X hours, ...

SLA's may refer to OLA's (Operational level agreements) [internal agreements of service within the division]

It's complicated and takes an enormous amount of work. But it's meant to prevent "This broke and we don't know who's responsible."

Muon g-2 will do a TSW at some point in the future

More yawning - sorry - on to good stuff...